



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Glorioso, Joseph C.  
Evans, Christopher H.  
Robbins, Paul D.
- (ii) TITLE OF INVENTION: Gene Transfer for Studying and Treating  
a Connective Tissue of a Mammalian Host
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Townsend and Townsend and Crew LLP
  - (B) STREET: Two Embarcadero Center, Eighth Floor
  - (C) CITY: San Francisco
  - (D) STATE: California
  - (E) COUNTRY: USA
  - (F) ZIP: 94111-3834
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/924,777
  - (B) FILING DATE: 05-SEP-1997
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/630,981
  - (B) FILING DATE: 20-DEC-1990
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 07/963,928
  - (B) FILING DATE: 20-OCT-1992
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/027,750
  - (B) FILING DATE: 08-MAR-1993
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/183,563
  - (B) FILING DATE: 18-JAN-1994
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/381,603
  - (B) FILING DATE: 27-JAN-1995
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/567,710
  - (B) FILING DATE: 05-DEC-1995
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 08/685,212
  - (B) FILING DATE: 23-JUL-1996

## (viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Bastian, Kevin L.
- (B) REGISTRATION NUMBER: 34,774
- (C) REFERENCE/DOCKET NUMBER: 018484-002280US

## (ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (415) 576-0200
- (B) TELEFAX: (415) 576-0300

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1770 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 55..1764
- (D) OTHER INFORMATION: /product= "human interleukin-1 receptor"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

CCTCCTGAGA AGCTGGACCC CTTGGTAAAA GACAAGGCCT TCTCCAAGAA GAAT ATG      57
                                     Met
                                     1

AAA GTG TTA CTC AGA CTT ATT TGT TTC ATA GCT CTA CTG ATT TCT TCT      105
Lys Val Leu Leu Arg Leu Ile Cys Phe Ile Ala Leu Leu Ile Ser Ser
                    5                      10                      15

CTG GAG GCT GAT AAA TGC AAG GAA CGT GAA GAA AAA ATA ATT TTA GTG      153
Leu Glu Ala Asp Lys Cys Lys Glu Arg Glu Glu Lys Ile Ile Leu Val
                20                      25                      30

TCA TCT GCA AAT GAA ATT GAT GTT CGT CCC TGT CCT CTT AAC CCA AAT      201
Ser Ser Ala Asn Glu Ile Asp Val Arg Pro Cys Pro Leu Asn Pro Asn
                35                      40                      45

GAA CAC AAA GGC ACT ATA ACT TGG TAT AAA GAT GAC AGC AAG ACA CCT      249
Glu His Lys Gly Thr Ile Thr Trp Tyr Lys Asp Asp Ser Lys Thr Pro
                50                      55                      60                      65

GTA TCT ACA GAA CAA GCC TCC AGG ATT CAT CAA CAC AAA GAG AAA CTT      297
Val Ser Thr Glu Gln Ala Ser Arg Ile His Gln His Lys Glu Lys Leu
                    70                      75                      80

TGG TTT GTT CCT GCT AAG GTG GAG GAT TCA GGA CAT TAC TAT TGC GTG      345
Trp Phe Val Pro Ala Lys Val Glu Asp Ser Gly His Tyr Tyr Cys Val
                    85                      90                      95

GTA AGA AAT TCA TCT TAC TGC CTC AGA ATT AAA ATA AGT GCA AAA TTT      393
Val Arg Asn Ser Ser Tyr Cys Leu Arg Ile Lys Ile Ser Ala Lys Phe
                100                      105                      110

```

GTG	GAG	AAT	GAG	CCT	AAC	TTA	TGT	TAT	AAT	GCA	CAA	GCC	ATA	TTT	AAG	441
Val	Glu	Asn	Glu	Pro	Asn	Leu	Cys	Tyr	Asn	Ala	Gln	Ala	Ile	Phe	Lys	
115						120					125					
CAG	AAA	CTA	CCC	GTT	GCA	GGA	GAC	GGA	GGA	CTT	GTG	TGC	CCT	TAT	ATG	489
Gln	Lys	Leu	Pro	Val	Ala	Gly	Asp	Gly	Gly	Leu	Val	Cys	Pro	Tyr	Met	
130					135					140					145	
GAG	TTT	TTT	AAA	AAT	GAA	AAT	AAT	GAG	TTA	CCT	AAA	TTA	CAG	TGG	TAT	537
Glu	Phe	Phe	Lys	Asn	Glu	Asn	Asn	Glu	Leu	Pro	Lys	Leu	Gln	Trp	Tyr	
				150					155					160		
AAG	GAT	TGC	AAA	CCT	CTA	CTT	CTT	GAC	AAT	ATA	CAC	TTT	AGT	GGA	GTC	585
Lys	Asp	Cys	Lys	Pro	Leu	Leu	Leu	Asp	Asn	Ile	His	Phe	Ser	Gly	Val	
			165					170						175		
AAA	GAT	AGG	CTC	ATC	GTG	ATG	AAT	GTG	GCT	GAA	AAG	CAT	AGA	GGG	AAC	633
Lys	Asp	Arg	Leu	Ile	Val	Met	Asn	Val	Ala	Glu	Lys	His	Arg	Gly	Asn	
		180					185					190				
TAT	ACT	TGT	CAT	GCA	TCC	TAC	ACA	TAC	TTG	GGC	AAG	CAA	TAT	CCT	ATT	681
Tyr	Thr	Cys	His	Ala	Ser	Tyr	Thr	Tyr	Leu	Gly	Lys	Gln	Tyr	Pro	Ile	
	195					200						205				
ACC	CGG	GTA	ATA	GAA	TTT	ATT	ACT	CTA	GAG	GAA	AAC	AAA	CCC	ACA	AGG	729
Thr	Arg	Val	Ile	Glu	Phe	Ile	Thr	Leu	Glu	Glu	Asn	Lys	Pro	Thr	Arg	
210					215					220					225	
CCT	GTG	ATT	GTG	AGC	CCA	GCT	AAT	GAG	ACA	ATG	GAA	GTA	GAC	TTG	GGA	777
Pro	Val	Ile	Val	Ser	Pro	Ala	Asn	Glu	Thr	Met	Glu	Val	Asp	Leu	Gly	
				230					235					240		
TCC	CAG	ATA	CAA	TTG	ATC	TGT	AAT	GTC	ACC	GGC	CAG	TTG	AGT	GAC	ATT	825
Ser	Gln	Ile	Gln	Leu	Ile	Cys	Asn	Val	Thr	Gly	Gln	Leu	Ser	Asp	Ile	
			245					250						255		
GCT	TAC	TGG	AAG	TGG	AAT	GGG	TCA	GTA	ATT	GAT	GAA	GAT	GAC	CCA	GTG	873
Ala	Tyr	Trp	Lys	Trp	Asn	Gly	Ser	Val	Ile	Asp	Glu	Asp	Asp	Pro	Val	
		260				265						270				
CTA	GGG	GAA	GAC	TAT	TAC	AGT	GTG	GAA	AAT	CCT	GCA	AAC	AAA	AGA	AGG	921
Leu	Gly	Glu	Asp	Tyr	Tyr	Ser	Val	Glu	Asn	Pro	Ala	Asn	Lys	Arg	Arg	
	275					280					285					
AGT	ACC	CTC	ATC	ACA	GTG	CTT	AAT	ATA	TCG	GAA	ATT	GAA	AGT	AGA	TTT	969
Ser	Thr	Leu	Ile	Thr	Val	Leu	Asn	Ile	Ser	Glu	Ile	Glu	Ser	Arg	Phe	
290					295				300						305	
TAT	AAA	CAT	CCA	TTT	ACC	TGT	TTT	GCC	AAG	AAT	ACA	CAT	GGT	ATA	GAT	1017
Tyr	Lys	His	Pro	Phe	Thr	Cys	Phe	Ala	Lys	Asn	Thr	His	Gly	Ile	Asp	
				310					315					320		
GCA	GCA	TAT	ATC	CAG	TTA	ATA	TAT	CCA	GTC	ACT	AAT	TTC	CAG	AAG	CAC	1065
Ala	Ala	Tyr	Ile	Gln	Leu	Ile	Tyr	Pro	Val	Thr	Asn	Phe	Gln	Lys	His	
			325					330					335			
ATG	ATT	GGT	ATA	TGT	GTC	ACG	TTG	ACA	GTC	ATA	ATT	GTG	TGT	TCT	GTT	1113
Met	Ile	Gly	Ile	Cys	Val	Thr	Leu	Thr	Val	Ile	Ile	Val	Cys	Ser	Val	
		340					345					350				

TTC	ATC	TAT	AAA	ATC	TTC	AAG	ATT	GAC	ATT	GTG	CTT	TGG	TAC	AGG	GAT	1161
Phe	Ile	Tyr	Lys	Ile	Phe	Lys	Ile	Asp	Ile	Val	Leu	Trp	Tyr	Arg	Asp	
355			360			365										
TCC	TGC	TAT	GAT	TTT	CTC	CCA	ATA	AAA	GCT	TCA	GAT	GGA	AAG	ACC	TAT	1209
Ser	Cys	Tyr	Asp	Phe	Leu	Pro	Ile	Lys	Ala	Ser	Asp	Gly	Lys	Thr	Tyr	
370			375			380			385							
GAC	GCA	TAT	ATA	CTG	TAT	CCA	AAG	ACT	GTT	GGG	GAA	GGG	TCT	ACC	TCT	1257
Asp	Ala	Tyr	Ile	Leu	Tyr	Pro	Lys	Thr	Val	Gly	Glu	Gly	Ser	Thr	Ser	
390			395			400										
GAC	TGT	GAT	ATT	TTT	GTG	TTT	AAA	GTC	TTG	CCT	GAG	GTC	TTG	GAA	AAA	1305
Asp	Cys	Asp	Ile	Phe	Val	Phe	Lys	Val	Leu	Pro	Glu	Val	Leu	Glu	Lys	
405			410			415										
CAG	TGT	GGA	TAT	AAG	CTG	TTC	ATT	TAT	GGA	AGG	GAT	GAC	TAC	GTT	GGG	1353
Gln	Cys	Gly	Tyr	Lys	Leu	Phe	Ile	Tyr	Gly	Arg	Asp	Asp	Tyr	Val	Gly	
420			425			430										
GAA	GAC	ATT	GTT	GAG	GTC	ATT	AAT	GAA	AAC	GTA	AAG	AAA	AGC	AGA	AGA	1401
Glu	Asp	Ile	Val	Glu	Val	Ile	Asn	Glu	Asn	Val	Lys	Lys	Ser	Arg	Arg	
435			440			445										
CTG	ATT	ATC	ATT	TTA	GTC	AGA	GAA	ACA	TCA	GGC	TTC	AGC	TGG	CTG	GGT	1449
Leu	Ile	Ile	Ile	Leu	Val	Arg	Glu	Thr	Ser	Gly	Phe	Ser	Trp	Leu	Gly	
450			455			460			465							
GGT	TCA	TCT	GAA	GAG	CAA	ATA	GCC	ATG	TAT	AAT	GCT	CTT	GTT	CAG	GAT	1497
Gly	Ser	Ser	Glu	Glu	Gln	Ile	Ala	Met	Tyr	Asn	Ala	Leu	Val	Gln	Asp	
470			475			480										
GGA	ATT	AAA	GTT	GTC	CTG	CTT	GAG	CTG	GAG	AAA	ATC	CAA	GAC	TAT	GAG	1545
Gly	Ile	Lys	Val	Val	Leu	Leu	Glu	Leu	Glu	Lys	Ile	Gln	Asp	Tyr	Glu	
485			490			495										
AAA	ATG	CCA	GAA	TCG	ATT	AAA	TTC	ATT	AAG	CAG	AAA	CAT	GGG	GCT	ATC	1593
Lys	Met	Pro	Glu	Ser	Ile	Lys	Phe	Ile	Lys	Gln	Lys	His	Gly	Ala	Ile	
500			505			510										
CGC	TGG	TCA	GGG	GAC	TTT	ACA	CAG	GGA	CCA	CAG	TCT	GCA	AAG	ACA	AGG	1641
Arg	Trp	Ser	Gly	Asp	Phe	Thr	Gln	Gly	Pro	Gln	Ser	Ala	Lys	Thr	Arg	
515			520			525										
TTC	TGG	AAG	AAT	GTC	AGG	TAC	CAC	ATG	CCA	GTC	CAG	CGA	CGG	TCA	CCT	1689
Phe	Trp	Lys	Asn	Val	Arg	Tyr	His	Met	Pro	Val	Gln	Arg	Arg	Ser	Pro	
530			535			540			545							
TCA	TCT	AAA	CAC	CAG	TTA	CTG	TCA	CCA	GCC	ACT	AAG	GAG	AAA	CTG	CAA	1737
Ser	Ser	Lys	His	Gln	Leu	Leu	Ser	Pro	Ala	Thr	Lys	Glu	Lys	Leu	Gln	
550			555			560										
AGA	GAG	GCT	CAC	GTG	CCT	CTC	GGG	TAGCATGGA								1770
Arg	Glu	Ala	His	Val	Pro	Leu	Gly									
565																

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 569 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Lys Val Leu Leu Arg Leu Ile Cys Phe Ile Ala Leu Leu Ile Ser
 1             5             10             15
Ser Leu Glu Ala Asp Lys Cys Lys Glu Arg Glu Glu Lys Ile Ile Leu
      20             25             30
Val Ser Ser Ala Asn Glu Ile Asp Val Arg Pro Cys Pro Leu Asn Pro
      35             40             45
Asn Glu His Lys Gly Thr Ile Thr Trp Tyr Lys Asp Asp Ser Lys Thr
      50             55             60
Pro Val Ser Thr Glu Gln Ala Ser Arg Ile His Gln His Lys Glu Lys
      65             70             75             80
Leu Trp Phe Val Pro Ala Lys Val Glu Asp Ser Gly His Tyr Tyr Cys
      85             90             95
Val Val Arg Asn Ser Ser Tyr Cys Leu Arg Ile Lys Ile Ser Ala Lys
      100            105            110
Phe Val Glu Asn Glu Pro Asn Leu Cys Tyr Asn Ala Gln Ala Ile Phe
      115            120            125
Lys Gln Lys Leu Pro Val Ala Gly Asp Gly Gly Leu Val Cys Pro Tyr
      130            135            140
Met Glu Phe Phe Lys Asn Glu Asn Asn Glu Leu Pro Lys Leu Gln Trp
      145            150            155            160
Tyr Lys Asp Cys Lys Pro Leu Leu Leu Asp Asn Ile His Phe Ser Gly
      165            170            175
Val Lys Asp Arg Leu Ile Val Met Asn Val Ala Glu Lys His Arg Gly
      180            185            190
Asn Tyr Thr Cys His Ala Ser Tyr Thr Tyr Leu Gly Lys Gln Tyr Pro
      195            200            205
Ile Thr Arg Val Ile Glu Phe Ile Thr Leu Glu Glu Asn Lys Pro Thr
      210            215            220
Arg Pro Val Ile Val Ser Pro Ala Asn Glu Thr Met Glu Val Asp Leu
      225            230            235            240
Gly Ser Gln Ile Gln Leu Ile Cys Asn Val Thr Gly Gln Leu Ser Asp
      245            250            255
Ile Ala Tyr Trp Lys Trp Asn Gly Ser Val Ile Asp Glu Asp Asp Pro
      260            265            270

```

Val Leu Gly Glu Asp Tyr Tyr Ser Val Glu Asn Pro Ala Asn Lys Arg  
 275 280 285  
 Arg Ser Thr Leu Ile Thr Val Leu Asn Ile Ser Glu Ile Glu Ser Arg  
 290 295 300  
 Phe Tyr Lys His Pro Phe Thr Cys Phe Ala Lys Asn Thr His Gly Ile  
 305 310 315 320  
 Asp Ala Ala Tyr Ile Gln Leu Ile Tyr Pro Val Thr Asn Phe Gln Lys  
 325 330 335  
 His Met Ile Gly Ile Cys Val Thr Leu Thr Val Ile Ile Val Cys Ser  
 340 345 350  
 Val Phe Ile Tyr Lys Ile Phe Lys Ile Asp Ile Val Leu Trp Tyr Arg  
 355 360 365  
 Asp Ser Cys Tyr Asp Phe Leu Pro Ile Lys Ala Ser Asp Gly Lys Thr  
 370 375 380  
 Tyr Asp Ala Tyr Ile Leu Tyr Pro Lys Thr Val Gly Glu Gly Ser Thr  
 385 390 395 400  
 Ser Asp Cys Asp Ile Phe Val Phe Lys Val Leu Pro Glu Val Leu Glu  
 405 410 415  
 Lys Gln Cys Gly Tyr Lys Leu Phe Ile Tyr Gly Arg Asp Asp Tyr Val  
 420 425 430  
 Gly Glu Asp Ile Val Glu Val Ile Asn Glu Asn Val Lys Lys Ser Arg  
 435 440 445  
 Arg Leu Ile Ile Ile Leu Val Arg Glu Thr Ser Gly Phe Ser Trp Leu  
 450 455 460  
 Gly Gly Ser Ser Glu Glu Gln Ile Ala Met Tyr Asn Ala Leu Val Gln  
 465 470 475 480  
 Asp Gly Ile Lys Val Val Leu Leu Glu Leu Glu Lys Ile Gln Asp Tyr  
 485 490 495  
 Glu Lys Met Pro Glu Ser Ile Lys Phe Ile Lys Gln Lys His Gly Ala  
 500 505 510  
 Ile Arg Trp Ser Gly Asp Phe Thr Gln Gly Pro Gln Ser Ala Lys Thr  
 515 520 525  
 Arg Phe Trp Lys Asn Val Arg Tyr His Met Pro Val Gln Arg Arg Ser  
 530 535 540  
 Pro Ser Ser Lys His Gln Leu Leu Ser Pro Ala Thr Lys Glu Lys Leu  
 545 550 555 560  
 Gln Arg Glu Ala His Val Pro Leu Gly  
 565

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1782 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 46..1776
- (D) OTHER INFORMATION: /product= "mouse interleukin-1 receptor"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGATGTCATC AGAGTTCCCA GTGCCCCGAA CCGTGAACAA CACAA ATG GAG AAT	54
Met Glu Asn	
1	
ATG AAA GTG CTA CTG GGG CTC ATT TGT CTC ATG GTG CCT CTG CTG TCG	102
Met Lys Val Leu Leu Gly Leu Ile Cys Leu Met Val Pro Leu Leu Ser	
5 10 15	
CTG GAG ATT GAC GTA TGT ACA GAA TAT CCA AAT CAG ATC GTT TTG TTT	150
Leu Glu Ile Asp Val Cys Thr Glu Tyr Pro Asn Gln Ile Val Leu Phe	
20 25 30 35	
TTA TCT GTA AAT GAA ATT GAT ATT CGC AAG TGT CCT CTT ACT CCA AAT	198
Leu Ser Val Asn Glu Ile Asp Ile Arg Lys Cys Pro Leu Thr Pro Asn	
40 45 50	
AAA ATG CAC GGC GAC ACC ATA ATT TGG TAC AAG AAT GAC AGC AAG ACC	246
Lys Met His Gly Asp Thr Ile Ile Trp Tyr Lys Asn Asp Ser Lys Thr	
55 60 65	
CCC ATA TCA GCG GAC CGG GAC TCC AGG ATT CAT CAG CAG AAT GAA CAT	294
Pro Ile Ser Ala Asp Arg Asp Ser Arg Ile His Gln Gln Asn Glu His	
70 75 80	
CTT TGG TTT GTA CCT GCC AAG GTG GAG GAC TCA GGA TAT TAC TAT TGT	342
Leu Trp Phe Val Pro Ala Lys Val Glu Asp Ser Gly Tyr Tyr Tyr Cys	
85 90 95	
ATA GTA AGA AAC TCA ACT TAC TGC CTC AAA ACT AAA GTA ACC GTA ACT	390
Ile Val Arg Asn Ser Thr Tyr Cys Leu Lys Thr Lys Val Thr Val Thr	
100 105 110 115	
GTG TTA GAG AAT GAC CCT GGC TTG TGT TAC AGC ACA CAG GCC ACC TTC	438
Val Leu Glu Asn Asp Pro Gly Leu Cys Tyr Ser Thr Gln Ala Thr Phe	
120 125 130	
CCA CAG CGG CTC CAC ATT GCC GGG GAT GGA AGT CTT GTG TGC CCT TAT	486
Pro Gln Arg Leu His Ile Ala Gly Asp Gly Ser Leu Val Cys Pro Tyr	
135 140 145	

GTG	AGT	TAT	TTT	AAA	GAT	GAA	AAT	AAT	GAG	TTA	CCC	GAG	GTC	CAG	TGG	534
Val	Ser	Tyr	Phe	Lys	Asp	Glu	Asn	Asn	Glu	Leu	Pro	Glu	Val	Gln	Trp	
		150					155					160				
TAT	AAG	AAC	TGT	AAA	CCT	CTG	CTT	CTT	GAC	AAC	GTG	AGC	TTC	TTC	GGA	582
Tyr	Lys	Asn	Cys	Lys	Pro	Leu	Leu	Leu	Asp	Asn	Val	Ser	Phe	Phe	Gly	
	165					170					175					
GTA	AAA	GAT	AAA	CTG	TTG	GTG	AGG	AAT	GTG	GCT	GAA	GAG	CAC	AGA	GGG	630
Val	Lys	Asp	Lys	Leu	Leu	Val	Arg	Asn	Val	Ala	Glu	Glu	His	Arg	Gly	
180					185					190					195	
GAC	TAT	ATA	TGC	CGT	ATG	TCC	TAT	ACG	TTC	CGG	GGG	AAG	CAA	TAT	CCG	678
Asp	Tyr	Ile	Cys	Arg	Met	Ser	Tyr	Thr	Phe	Arg	Gly	Lys	Gln	Tyr	Pro	
			200						205					210		
GTC	ACA	CGA	GTA	ATA	CAA	TTT	ATC	ACA	ATA	GAT	GAA	AAC	AAG	AGG	GAC	726
Val	Thr	Arg	Val	Ile	Gln	Phe	Ile	Thr	Ile	Asp	Glu	Asn	Lys	Arg	Asp	
			215					220					225			
AGA	CCT	GTT	ATC	CTG	AGC	CCT	CGG	AAT	GAG	ACG	ATC	GAA	GCT	GAC	CCA	774
Arg	Pro	Val	Ile	Leu	Ser	Pro	Arg	Asn	Glu	Thr	Ile	Glu	Ala	Asp	Pro	
	230						235					240				
GGA	TCA	ATG	ATA	CAA	CTG	ATC	TGC	AAC	GTC	ACG	GGC	CAG	TTC	TCA	GAC	822
Gly	Ser	Met	Ile	Gln	Leu	Ile	Cys	Asn	Val	Thr	Gly	Gln	Phe	Ser	Asp	
	245					250					255					
CTT	GTC	TAC	TGG	AAG	TGG	AAT	GGA	TCA	GAA	ATT	GAA	TGG	AAT	GAT	CCA	870
Leu	Val	Tyr	Trp	Lys	Trp	Asn	Gly	Ser	Glu	Ile	Glu	Trp	Asn	Asp	Pro	
260					265					270					275	
TTT	CTA	GCT	GAA	GAC	TAT	CAA	TTT	GTG	GAA	CAT	CCT	TCA	ACC	AAA	AGA	918
Phe	Leu	Ala	Glu	Asp	Tyr	Gln	Phe	Val	Glu	His	Pro	Ser	Thr	Lys	Arg	
				280					285					290		
AAA	TAC	ACA	CTC	ATT	ACA	ACA	CTT	AAC	ATT	TCA	GAA	GTT	AAA	AGC	CAG	966
Lys	Tyr	Thr	Leu	Ile	Thr	Thr	Leu	Asn	Ile	Ser	Glu	Val	Lys	Ser	Gln	
		295					300						305			
TTT	TAT	CGC	TAT	CCG	TTT	ATC	TGT	GTT	GTT	AAG	AAC	ACA	AAT	ATT	TTT	1014
Phe	Tyr	Arg	Tyr	Pro	Phe	Ile	Cys	Val	Val	Lys	Asn	Thr	Asn	Ile	Phe	
	310					315						320				
GAG	TCG	GCG	CAT	GTG	CAG	TTA	ATA	TAC	CCA	GTC	CCT	GAC	TTC	AAG	AAT	1062
Glu	Ser	Ala	His	Val	Gln	Leu	Ile	Tyr	Pro	Val	Pro	Asp	Phe	Lys	Asn	
	325					330					335					
TAC	CTC	ATC	GGG	GGC	TTT	ATC	ATC	CTC	ACG	GCT	ACA	ATT	GTA	TGC	TGT	1110
Tyr	Leu	Ile	Gly	Gly	Phe	Ile	Ile	Leu	Thr	Ala	Thr	Ile	Val	Cys	Cys	
340					345					350					355	
GTG	TGC	ATC	TAT	AAA	GTC	TTC	AAG	GTT	GAC	ATA	GTG	CTT	TGG	TAC	AGG	1158
Val	Cys	Ile	Tyr	Lys	Val	Phe	Lys	Val	Asp	Ile	Val	Leu	Trp	Tyr	Arg	
				360					365					370		
GAC	TCC	TGC	TCT	GGT	TTT	CTT	CCT	TCA	AAA	GCT	TCA	GAT	GGA	AAG	ACA	1206
Asp	Ser	Cys	Ser	Gly	Phe	Leu	Pro	Ser	Lys	Ala	Ser	Asp	Gly	Lys	Thr	
			375					380					385			



TAC	GAT	GCA	TAT	ATT	CTT	TAT	CCC	AAG	ACC	CTG	GGA	GAG	GGG	TCC	TTC	1254
Tyr	Asp	Ala	Tyr	Ile	Leu	Tyr	Pro	Lys	Thr	Leu	Gly	Glu	Gly	Ser	Phe	
		390					395					400				
TCA	GAC	TTA	GAT	ACT	TTT	GTT	TTT	AAA	CTG	TTG	CCT	GAG	GTC	TTG	GAG	1302
Ser	Asp	Leu	Asp	Thr	Phe	Val	Phe	Lys	Leu	Leu	Pro	Glu	Val	Leu	Glu	
		405				410					415					
GGA	CAG	TTT	GGA	TAC	AAG	CTG	TTC	ATT	TAT	GGA	AGG	GAT	GAC	TAT	GTT	1350
Gly	Gln	Phe	Gly	Tyr	Lys	Leu	Phe	Ile	Tyr	Gly	Arg	Asp	Asp	Tyr	Val	
		420			425					430					435	
GGA	GAA	GAT	ACC	ATC	GAG	GTT	ACT	AAT	GAA	AAT	GTA	AAG	AAA	AGC	AGG	1398
Gly	Glu	Asp	Thr	Ile	Glu	Val	Thr	Asn	Glu	Asn	Val	Lys	Lys	Ser	Arg	
				440					445						450	
AGG	CTG	ATT	ATC	ATT	CTA	GTG	AGA	GAT	ATG	GGA	GGC	TTC	AGC	TGG	CTG	1446
Arg	Leu	Ile	Ile	Ile	Leu	Val	Arg	Asp	Met	Gly	Gly	Phe	Ser	Trp	Leu	
			455					460					465			
GGC	CAG	TCA	TCT	GAA	GAG	CAA	ATA	GCC	ATA	TAC	AAT	GCT	CTC	ATC	CAG	1494
Gly	Gln	Ser	Ser	Glu	Glu	Gln	Ile	Ala	Ile	Tyr	Asn	Ala	Leu	Ile	Gln	
		470					475					480				
GAA	GGA	ATT	AAA	ATC	GTC	CTG	CTT	GAG	TTG	GAG	AAA	ATC	CAA	GAC	TAT	1542
Glu	Gly	Ile	Lys	Ile	Val	Leu	Leu	Glu	Leu	Glu	Lys	Ile	Gln	Asp	Tyr	
		485				490					495					
GAG	AAA	ATG	CCA	GAT	TCT	ATT	CAG	TTC	ATT	AAG	CAG	AAA	CAC	GGA	GTC	1590
Glu	Lys	Met	Pro	Asp	Ser	Ile	Gln	Phe	Ile	Lys	Gln	Lys	His	Gly	Val	
		500			505					510					515	
ATT	TGC	TGG	TCA	GGA	GAC	TTT	CAA	GAA	AGA	CCA	CAG	TCT	GCA	AAG	ACC	1638
Ile	Cys	Trp	Ser	Gly	Asp	Phe	Gln	Glu	Arg	Pro	Gln	Ser	Ala	Lys	Thr	
				520					525					530		
AGG	TTC	TGG	AAA	AAC	TTA	AGA	TAC	CAG	ATG	CCA	GCC	CAA	CGG	AGA	TCA	1686
Arg	Phe	Trp	Lys	Asn	Leu	Arg	Tyr	Gln	Met	Pro	Ala	Gln	Arg	Arg	Ser	
			535					540					545			
CCA	TTG	TCT	AAA	CAC	CGC	TTA	CTA	ACC	CTG	GAT	CCT	GTG	CGG	GAC	ACT	1734
Pro	Leu	Ser	Lys	His	Arg	Leu	Leu	Thr	Leu	Asp	Pro	Val	Arg	Asp	Thr	
			550				555					560				
AAG	GAG	AAA	CTG	CCG	GCA	GCA	ACA	CAC	TTA	CCA	CTC	GGC	TAGCATGGC			1782
Lys	Glu	Lys	Leu	Pro	Ala	Ala	Thr	His	Leu	Pro	Leu	Gly				
		565				570					575					

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 576 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Glu	Asn	Met	Lys	Val	Leu	Leu	Gly	Leu	Ile	Cys	Leu	Met	Val	Pro	1	5	10	15
Leu	Leu	Ser	Leu	Glu	Ile	Asp	Val	Cys	Thr	Glu	Tyr	Pro	Asn	Gln	Ile	20	25	30	
Val	Leu	Phe	Leu	Ser	Val	Asn	Glu	Ile	Asp	Ile	Arg	Lys	Cys	Pro	Leu	35	40	45	
Thr	Pro	Asn	Lys	Met	His	Gly	Asp	Thr	Ile	Ile	Trp	Tyr	Lys	Asn	Asp	50	55	60	
Ser	Lys	Thr	Pro	Ile	Ser	Ala	Asp	Arg	Asp	Ser	Arg	Ile	His	Gln	Gln	65	70	75	80
Asn	Glu	His	Leu	Trp	Phe	Val	Pro	Ala	Lys	Val	Glu	Asp	Ser	Gly	Tyr	85	90	95	
Tyr	Tyr	Cys	Ile	Val	Arg	Asn	Ser	Thr	Tyr	Cys	Leu	Lys	Thr	Lys	Val	100	105	110	
Thr	Val	Thr	Val	Leu	Glu	Asn	Asp	Pro	Gly	Leu	Cys	Tyr	Ser	Thr	Gln	115	120	125	
Ala	Thr	Phe	Pro	Gln	Arg	Leu	His	Ile	Ala	Gly	Asp	Gly	Ser	Leu	Val	130	135	140	
Cys	Pro	Tyr	Val	Ser	Tyr	Phe	Lys	Asp	Glu	Asn	Asn	Glu	Leu	Pro	Glu	145	150	155	160
Val	Gln	Trp	Tyr	Lys	Asn	Cys	Lys	Pro	Leu	Leu	Leu	Asp	Asn	Val	Ser	165	170	175	
Phe	Phe	Gly	Val	Lys	Asp	Lys	Leu	Leu	Val	Arg	Asn	Val	Ala	Glu	Glu	180	185	190	
His	Arg	Gly	Asp	Tyr	Ile	Cys	Arg	Met	Ser	Tyr	Thr	Phe	Arg	Gly	Lys	195	200	205	
Gln	Tyr	Pro	Val	Thr	Arg	Val	Ile	Gln	Phe	Ile	Thr	Ile	Asp	Glu	Asn	210	215	220	
Lys	Arg	Asp	Arg	Pro	Val	Ile	Leu	Ser	Pro	Arg	Asn	Glu	Thr	Ile	Glu	225	230	235	240
Ala	Asp	Pro	Gly	Ser	Met	Ile	Gln	Leu	Ile	Cys	Asn	Val	Thr	Gly	Gln	245	250	255	
Phe	Ser	Asp	Leu	Val	Tyr	Trp	Lys	Trp	Asn	Gly	Ser	Glu	Ile	Glu	Trp	260	265	270	
Asn	Asp	Pro	Phe	Leu	Ala	Glu	Asp	Tyr	Gln	Phe	Val	Glu	His	Pro	Ser	275	280	285	
Thr	Lys	Arg	Lys	Tyr	Thr	Leu	Ile	Thr	Thr	Leu	Asn	Ile	Ser	Glu	Val	290	295	300	
Lys	Ser	Gln	Phe	Tyr	Arg	Tyr	Pro	Phe	Ile	Cys	Val	Val	Lys	Asn	Thr	305	310	315	320

Asn	Ile	Phe	Glu	Ser	Ala	His	Val	Gln	Leu	Ile	Tyr	Pro	Val	Pro	Asp	325	330	335
Phe	Lys	Asn	Tyr	Leu	Ile	Gly	Gly	Phe	Ile	Ile	Leu	Thr	Ala	Thr	Ile	340	345	350
Val	Cys	Cys	Val	Cys	Ile	Tyr	Lys	Val	Phe	Lys	Val	Asp	Ile	Val	Leu	355	360	365
Trp	Tyr	Arg	Asp	Ser	Cys	Ser	Gly	Phe	Leu	Pro	Ser	Lys	Ala	Ser	Asp	370	375	380
Gly	Lys	Thr	Tyr	Asp	Ala	Tyr	Ile	Leu	Tyr	Pro	Lys	Thr	Leu	Gly	Glu	385	390	395
Gly	Ser	Phe	Ser	Asp	Leu	Asp	Thr	Phe	Val	Phe	Lys	Leu	Leu	Pro	Glu	405	410	415
Val	Leu	Glu	Gly	Gln	Phe	Gly	Tyr	Lys	Leu	Phe	Ile	Tyr	Gly	Arg	Asp	420	425	430
Asp	Tyr	Val	Gly	Glu	Asp	Thr	Ile	Glu	Val	Thr	Asn	Glu	Asn	Val	Lys	435	440	445
Lys	Ser	Arg	Arg	Leu	Ile	Ile	Ile	Leu	Val	Arg	Asp	Met	Gly	Gly	Phe	450	455	460
Ser	Trp	Leu	Gly	Gln	Ser	Ser	Glu	Glu	Gln	Ile	Ala	Ile	Tyr	Asn	Ala	465	470	475
Leu	Ile	Gln	Glu	Gly	Ile	Lys	Ile	Val	Leu	Leu	Glu	Leu	Glu	Lys	Ile	485	490	495
Gln	Asp	Tyr	Glu	Lys	Met	Pro	Asp	Ser	Ile	Gln	Phe	Ile	Lys	Gln	Lys	500	505	510
His	Gly	Val	Ile	Cys	Trp	Ser	Gly	Asp	Phe	Gln	Glu	Arg	Pro	Gln	Ser	515	520	525
Ala	Lys	Thr	Arg	Phe	Trp	Lys	Asn	Leu	Arg	Tyr	Gln	Met	Pro	Ala	Gln	530	535	540
Arg	Arg	Ser	Pro	Leu	Ser	Lys	His	Arg	Leu	Leu	Thr	Leu	Asp	Pro	Val	545	550	555
Arg	Asp	Thr	Lys	Glu	Lys	Leu	Pro	Ala	Ala	Thr	His	Leu	Pro	Leu	Gly	565	570	575

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCGGATCCCC TCCTGAGAAG CT

22

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 21 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCGGATCCCA TGTGCTACTG G

21

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 18 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGCACCATGG TACCTGCA

18

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 21 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: single  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCACAGGAT CCTCTGGGTA C

21